**Lab Virtual Screen 4: Validation dock against NDM-1 Metallo β-lactamase** REDO v.2

**Modifications for REDO:**

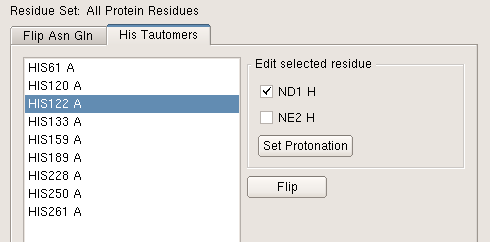
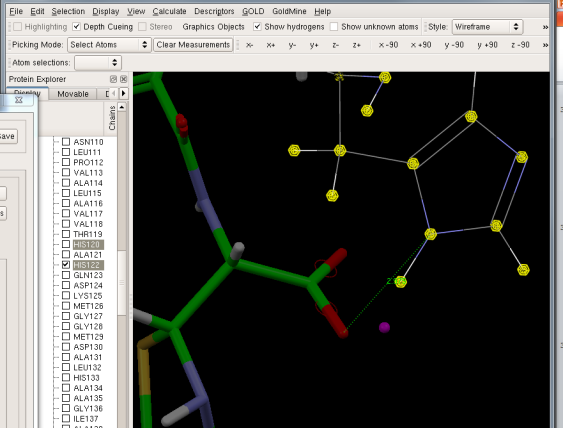
On the DDFE, make a copy of your VS4 protocol folder, call it VS4redo

Make a copy of your Molprobity prepared protein file (.pdb) which has FLIPS, but no H’s added.

* Rename to add ‘NOCONECTS’ to the end. (e.g. 4eybchAF.pdb 🡪 4eybchAFNOCONECTS.pdb)
* Delete the CONECT records to the metals (e.g. zinc) in your PDB file after it has gone through Molprobity (NOTE: this is tricky – you only want to take out the references to the metals connections to other atoms while preserving any other connections!)

Edit your configuration file so that the paths refer to the new folder now and the new protein file name.

Start Xmin, Xlaunch, putty, go into VS4redo folder  
Open GOLD GUI, however this time 'open' the configuration file that you just put into that folder instead of starting a new one  
Go to protonation states and apply proper protonation to Histidine residues in the active site  
(when you click on the name of the HIS e.g. ‘HIS120A’ it will be highlighted in GOLD Hermes.

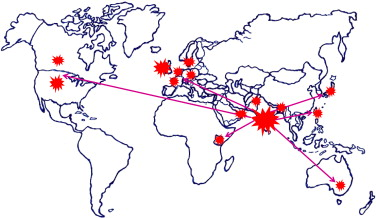
* + HINT: in the Hermes protein explorer panel on the left, you can hide all of the residues except the Histidines you are interested in and the metals. This will de-clutter the image and make it easier to determine protonation state.
* If a HIS is not near the active site, you don’t really care about it.
* To change the protonation state, toggle the ‘ND1 H’ and ‘NE2 H’ buttons. Then hit ‘Set Protonation’. Don’t worry you can change it back – try it out. Also, try ‘flipping’ the HIS.

OPTIONAL: You may want to open up a PyMol session of your crystal structure so that you can easily view the histidines (the Hermes interface can be 'bulky' and slow). This is also a good way to make an image for a poster or report.  
- open the original PDB in PyMol

* Find your metals, make them a selection, then show as spheres
* find ALL histidines in their by typing: $select histidines, resn his, show them all as sticks
* zoom in on active site  
  - change selecting mode to 'atoms'  
  - click on one atom each on each of the histidines in the active site (should be 2-4 of them)  
  - rename the selection 'activehistidines'  
  - Label those atoms using the 'L' tab for that selection - show Residue identifier (i.e. a number)  
  - use this image and the residue numbers to help you protonate the Histidines in the GOLD Hermes window

Re-run the docking using the same ligands  
Post results in table on Wikispaces NDM-1 Beta-lactamase Project alongside your original VS4 results

**ORIGINAL PROTOCOL:**

**OBJECTIVE:** The purpose of this lab is to validate GOLD as a docking resource against the Metallo β-lactamase class of enzymes.

**INTRODUCTION**

**Disease and Target:**

See this PDB feature on the NDM-1 target for background information:

<http://sbkb.org/kb/featured_molecule_latest.html>

**VIRTUAL SCREENING**

* One Run
  + Screen ~6 ligands at 2.0 autoscale on **1** processor
* Analyze Bestranking list, post scores to Wikispaces
* Examine poses in PyMol

In your lab notebook, record your steps and modificatons (or choices/options) that you make in the protocol.

**DOWNLOAD THE CRYSTAL STRUCTURE**:

Go to **PDB website**, do the search for ‘**NDM-1’**

Pick one of the bacterial versions that has one ligand already in the active site. It can have other ligands – as long as there is only one in the active site.

Select one that satisfies the following criteria:

* + Has only 1 inhibitor present in the active site (cannot be allosteric either)
    - NOTE: It is ok if it has some other ‘ligands’ in solution around the protein – just as long as they are not in the active site. Other molecules that may be in solution but are not really ‘ligand’ are: Sodium, Potassium, Glycerol, Tris.
    - TIP: Use PyMol or the built-in Jmol viewer to verify where a ligand is in the structure.
  + Choose a file that has no mutations in the amino acid sequence of the active site
    - It is ok to have mutations peripherally though
    - Best if it does not contain **Modified Residues** (i.e. MSE (selenomethionine))

Only **5** people can dock against each of the structures. Check which ones are taken on this website and then enter your name once you have chosen one.

<http://vdsstream.wikispaces.com/NDM-1+Beta-lactamase+Project>

**DOWNLOAD THE PROTEIN –**

Make note of:

**How many metal ions are present?**

**What type of metal is present?**

**Write down what organism you chose & the PDB ID**

**DOWNLOAD THE LIGAND**

Also, download the PubChem version of the ligand (don’t download the PDB version – it isn’t properly configured). In the PDB page, go to **Ligand Chemical Component**. Click on the Identifier for the ligand (e.g. PSY) and this will take you to the **Ligand Summary** page. Now, follow the PubChem link to download the structure as 3D by clicking on the ‘**SDF´**icon in the top right of the PubChem page. Choose to save as a **3D SDF.** Keep the name given.

NOTE: this ligand may not be fully minimized. We will actually skip the ligand minimization step for the sake of time/simplicity.

**VERIFY AND FIX STRUCTURES IN PYMOL**:

Open the ligand in PyMol and make sure it looks legit **Q:** is it in 3D? Does it have Hydrogens?

Now, open the protein file in PyMol. IF there is more than one chain present we will want to remove them. Open your structure in PyMol, and Color ‘**By Chain’** You should see each of the chains as a different color. Now, change your ‘**Selecting’** mode to **‘Chains’** click on the chain you want to get rid of (usually B, C, or D – but make sure the one you keep has the ligand present in the active site) . When it is made as a new selection by PyMol – **Remove Atoms**

**Then go to File >> Save Molecule >> as**

**<the pdb identifier>chainA.pdb for example: XXXXchainA.pdb or XXXXchainB.pdb**

**MOLPROBITY ANALYSIS OF PROTEIN**

Go to MolProbity website – TIP: use Internet Explorer

‘**Choose File**’ - Load **your PDB** file to the website (use the file you have from above instead of entering the PDB identifier directly into MolProbity)

Copy down the resolution (to put in your table below)

Add Hydrogens,

On next page, Choose the option to Add the flips ‘**Asn/Gln/His flips’**

Leave all the defaults checked and ‘**regenerate H’s’**

If it asks you to save the file with Hydrogens – skip it for now.

**Analyze all-atom contacts and geometry**

Use all defaults **>> Run Programs to Perform these Analyses**

**View the Multi-Criterion Chart**

Save the data for your table below

View the **Multi-Criterion Kinemage** to see what type of errors exist in the active site.

Make note of which type they are for your table below.

HINT: to see the active site – toggle on and off the **‘hets’** button

**Include** a ‘snip’ of your traffic light table from MolProbity (i.e. green, yellow, red table)

**Also include** these values for the structure in your notebook. You can get the resolution from the PDB site.

|  |
| --- |
| **Resolution:** |
| **Errors in Active Site:** |

**Q:** How good is the structure relative to the ones you used for the VS2 and VS3 lab?

**Download** your PDB file with REMARK 40 WITHOUT Hydrogens – we will make GOLD add the H’s later

The filename should show up as something like:

XXXXchainAF.pdb (not XXXXchainAFH.pdb)

**TRANSFER PROTEIN AND LIGAND TO DDFE**

Login to the DDFE using WinSCP

ddfe.cm.utexas.edu

In virtual screening, it is important to keep your file structure organized and to reduce redundant files.

Create a directory structure on the DDFE like this below.

e.g. /home/chem204/2013/YOURUTEID/**LabVS4\_NDM-1**

Transfer the **protein** (.pdb) - with flips but without hydrogens - from your local computer over to the DDFE.

**LIGAND LIBRARY SELECTION - POSITIVE AND NEGATIVE CONTROLS**

We need to get some positive and negative controls to make the library. You will need to gather the physico-chemical properties (**Molecular weight, H-bond Donors, H-bond Acceptors, LogP**) for your final results table and also download the SDF files.

1. **POSITIVE CONTROL #1**

Go to the BindingDB database (<http://www.bindingdb.org/bind/index.jsp>)

Under the **Full Search** type in: “Metallo-beta-lactamase”

Scroll down to Target: **metallo-beta-lactamase**

The first entry is a potential inhibitor for a similar beta-lactamase

Make note of the **Ki** value in **nM or IC50 value – if present**

Click on the **PC cid** and then download the structure as 3D by clicking on the ‘**SDF´**icon in the top right. Choose to save as **3D SDF** and record the **Physico-Chemical properties.**

* add the name ‘**pos1’** to the end of the file (e.g. CID\_ 44542081**pos1**.sdf)
* Open the file in WORDPAD (right click >> Open With) and add ‘**pos1’** to the end of the identifier in the 1st line

E.g. 44542081**pos1**

1. **POSITIVE CONTROL – faropenem**

Faropenem is a beta-lactam drug in the penem class. Get the structure from PubChem directly and record the **Physico-Chemical properties.**

* add the name ‘**faropenem’** to the end of the file & to the end of the 1st line as you did above

1. **POSITIVE CONTROL – penicillin**

Penicillin is the most common beta-lactam drug. Get the structure from PubChem directly and record the **Physico-Chemical properties.**

* add the name ‘**penicliin**’ to the end of the file & to the end of the 1st line as you did above

1. **POSITIVE CONTROL – clavulanic acid**

Clavulanic Acid is the most common inhibitor of standard beta-lactamases. Get the structure from PubChem and record the **Physico-Chemical properties.**

* add the name ‘**clavulanic acid** to the end of the file & to the end of the 1st line as you did above

1. **RANDOM CONTROL - aspririn**

Find **Aspririn** in the PubChem database and download it as **3D** **SDF** and record the **Physico-Chemical properties.**. We don’t have any evidence that aspirin will bind. So, it is a RANDOM or NEGATIVE CONTROL

* add the name ‘**aspirin’** to the end of the file & to the end of the 1st line as you did above

1. **Original Ligand that is in the PDB**

This is the ligand you downloaded on page 2 from PubChem.

Docking this is essentially a ‘validation dock’ to see if you docking program can properly dock a true ligand.

* add the name ‘**original’** to the end of the file & to the end of the 1st line as you did above

You should have 6 new compounds now from the PubChem database. Open each of them in PyMol to verify they look ‘ok’

**NOTE**: For the next two steps - you can concatenate all 6 ligand together first, then pH them in OpenBabel. Or pH them individually and then concatenate.

**LIGAND PROTONATION STATE (pH’ing)**

The ligands may or may not be in the correct protonation state when they are downloaded from the PubChem Database. Usually we will dock ligands at physiological pH – which is pH 7.4 (or 7.2).

There are 2 ways to do this

1. Go to OpenBabel on a Windows computer.

* Choose SDF as the INPUT FORMAT type, browse for your ligand file
* Enter **“7.4”** to ‘Add hydrogens appropriate for this pH’
* Output Format = SDF
* Name your output file in a useful manner. E.g. **CID\_378pNPP\_pH.sdf**
* **CONVERT**
* **Repeat this for each of your Controls**
* Open each of them in PyMol to verify they look ‘ok’ – do the hydrogens change from the original?
* Transfer to the DDFE folder

1. Use command line version of Babel

* Transfer your files to the DDFE
* Open a Terminal Window in the folder where your files are
* Type this command to add Hydrogens appropriate for pH 7.4

$babel –isdf inputfilename.sdf –osdf outputfilename.sdf –p 7.4

e.g. $babel –isdf CID\_378pNPP.sdf –osdf CID\_378pNPP\_ph.sdf –p 7.4

* Open each of them in PyMol to verify they look ‘ok’– do the hydrogens change from the original?

**CONCATENATING THE LIBRARY**

Now **append** all of these together.

$cat <FIRSTLIGAND>.sdf <SECONDLIGAND>.sdf <THIRDLIGAND>.sdf <FOURTHLIGAND>.sdf <FIFTHLIGAND>.sdf <SIXTHLIGAND>.sdf >> **ValidationLibraryAll\_pH.sdf**

Where you fill in the real filenames in place of the brackets <>. Your concatenated file should have 6 ligands

**CHECK YOUR LIGAND LIBRARY**

Transfer your newly made library over to your local computer. Open it in **PyMol** and check to be sure all the ligands with your Controls are present.

**IF,** your library failed to be assemble correctly, you may need to edit the ligand files.

Open the files in WORDPAD.

Check to see that there is only one empty line after the four dollar symbols ($$$$)

If there are two empty lines, remove the last empty line at the end of the file!

Otherwise, the files cannot be joined together properly!

**Q:** State an hypothesis about how you think these compounds will fare in the virtual screening of the target that you selected. Address the controls you are using and the novel ligands.

**PROTEIN PREPARATION**

**NOTE:** these instructions are for Windoze. If you insist upon using your MacIntrash – see the Tips&Hints Page of our Wikispaces page for help.

**Connecting to the graphical interface for GOLD**

Make remote connection to DDFE using a graphical user interface (GUI) for GOLD

Open Xming server

Go to Start, Programs, Xming, Xming

Open Xlaunch

Go to Start, Programs, Xming, XLaunch

Select ‘Multiple Windows’

Select ‘Start no client’,

Skip next screen by selecting ‘Next’ then ‘Finish’ on next screen

Open **Putty** in Programs

Connect to Host Name: ddfe.cm.utexas.edu on Port 22 using SSH

On the left side of the window, Select the ‘SSH’ tab and then the ‘X11’ or ‘Tunnels’ tab

‘Enable X11 forwarding’

X display location: leave blank or enter localhost:0

# this is the default display on your computer

‘Open’

Login as user: type your user name for the DDFE (your UTEID)

Enter password

Then type this to open gold with the graphical user interface:

**$gold**

Step through the Configuration Options to set up your file

Skip Wizard, Skip Templates

**Protein > Load protein** <nameofyourpdbfile>.pdb

Gold has a separate window for Global Options and a specific window for operations on your protein.

**Under the tab for your PDB file name (to the right) – the tab may just say “ID”:**

**Protonation & Tautomers** > Add Hydrogens - Write down how many hydrogens added.

Skip - **Flip Asn GLn and HIS tautomers**. We won’t worry about these right now.

**Extract/Delete Waters**: ‘Delete Remaining Waters’ (don’t select any of them to save).

Write down how many waters removed.

**Delete Ligand**

NOTE: If there is more than one ligand, you will need to go into the **Hermes** visualizer window to figure out which ligand

Go to View >> Protein Explorer

Click on the ‘+’ (plus sign) to see the different objects.

Extract the Ligand, save as ‘**LigandExtracted.mol2’**

(this will be saved for defining the cavity site)

NOTE: you already saved a different PubChem version of this ligand for validation docking

**Back in WinSCP** – make sure your LigandExtracted has an extension

If not, then add it to the file (just add **.mol2** to the end)

\* Go to **METALS** – and check that GOLD sees the right number (and type) of metal.

Leave the default metal coordination

Skip the remaining options for the protein.

**Under Global Options:**

Define Binding Site –‘Select One or more ligands’

‘One or more ligands’ - choose the single ligand that you had extracted.

‘Select all atoms within 7.5 Angstroms

Leave ‘Generate a cavity’ unchecked

Check – ‘Detect cavity’

Check – ‘Force all H bond donors/acceptors ….”

– verify active site in image on the Hermes visualizer

(only a small region around the ligand of the protein will be highlighted in gray)

In the Gold GUI – go back to **Global Options**

Select Ligands – you will need find where your library is (probably in your VS4 directory)

This is the file you need to link to for your ligand library.

Then make sure the number of conformations per ligand or **GA Runs** is set to ‘10’

Skip the Reference Ligand

Skip ‘Configure Waters’

Skip ‘Ligand Flexibility’

Leave the defaults for ‘Fitness & Search Options’ - it will use CHEMPLP scoring function.

‘GA Settings’ – 200%

Output Options

Output directory: leave as it is (‘.’)

UNCHECK – save ligand rank (.rnk) files

UNCHECK – save ligand log files

UNCHECK – save initialized ligand files

Save solutions to one file:

‘YourTargetvsYourLibrary.sdf’ e.g. “<YOURPROTEINNAME>vsValidationLibrary.sdf”

bestranking\_list\_filename

‘BestYourTargetvsYourLibrary.lst’ e.g. “Best<YOURPROTEINNAME>vsValidationLibrary.lst”

Skip ‘Information in File’

Under the ‘Selecting Solutions’ tab – select

Keep only the **single** best ranked solution (pose) for each of the ligands.

Skip GoldMine

Skip Parallel GOLD – we will run in parallel but it will be executed remotely instead of at this console

Skip ‘Constraints’

‘Atom Typing’ - Automatically **set atom and bond types (for the ligand only)**:

Make sure only one box is checked - ‘Ligand’ only

At the top of the page hit Save

Hit ‘Finish’ to save the file

Save GOLD conf file as **gold.conf**

Save protein as <**PDBname>protein.mol2**

Then close GOLD/Hermes

**VERIFY GOLD.CONF**

go back to WinSCP to VERIFY your newly made gold.conf file and MODIFY it in Wordpad

Set Autoscale to **2.0**

cavity\_file = Cavity file name that you made in the Hermes prep - YourLigand.mol2

-may need to add the extension manually if not present

(do this on the actual ligand file and on the line in gold.conf file)

ligand\_data\_file = Reference to ligand file set (the whole path needs to be there).

Number of conformers is ‘10’

Be sure that ‘set\_protein\_atom\_types = 0’

directory = .

concatenated\_output = verify it matches what you entered before

clean\_up\_option save\_best\_ligands = should be all of them

protein\_datafile = Protein target file name (PDBname\_protein.mol2 – from Hermes prep)

**JOB SCRIPT**

Obtain a **scriptgoldscanshisjob.sh** from the **/home/chem204/scripts** directory and modify it to run for this job.

**TIP**: it helps to have your **gold.conf** file open in a separate window at the same time to look at.

To say:

#!/bin/csh

$GOLD\_DIR/gold\_scan <firstligand> <lastligand> <slice> <PATHTOLIBRARY>.sdf

We will run on 1 processor.

Verify they are running by using the command to check the status of the queue.

**$qstat**

**TIP:** if it fails and you need to delete a whole directory – you need to type this to remove it recursively (i.e. with all files inside)

**$rm –r directoryname**

If it does not run, Save the gold.err file and show to a Mentor or Dr. B

**RUN ANALYSIS**

**VERIFY POSES**

Transfer this new output SDF file to the Desktop. To check that these output poses actually dock in to the active site of the protein, open the original PDB file in PyMol along with these output ligand poses from your Run1.

**Q:** Do the ligands dock in the active site?

**Q:** Are there the right number of ligands in the Output?

**CHECKING FITNESS SCORES**

Enter the bestranking values onto the Wikispaces page to compare amongst the class.

<http://vdsstream.wikispaces.com/NDM-1+Beta-lactamase+Project>

**Make a table of the ligands for your notebook**

In Excel, use the Text Import Wizard. Open Excel first, and then Open the files (remember to force Excel to ‘All Files’).

**Sort by Fitness score**.

Add a column called ‘**RANK’** and put the rank of the compound.

Fix the header row along the ‘**Fitness’** line so that the column labels read correctly.

Also, adjust the column width to fit the data.

The GOLD output should show S(PLP) hydrophobic score, S(hbond) hydrogen bond score, S(cho), S(metal), DE(clash), DE(tors), and ‘intcor’ and the ligand identification (a 6-7 digit number). You DO NOT need the File name, or Time shown in your table.

**Give your table a super-awesome caption!**

**Lipinski’s Data**

Include a 2-D image of the compounds and the associated Lipinski’s data. You can find this information about the compounds - LogP, number of hydrogen bond donors and acceptors, and the molecular weights – from the PubChem database. <http://pubchem.ncbi.nlm.nih.gov/>

The **2-D pictures** can be pasted separately from the table but they need to be labeled by compound name/identifier to easily identify which is which.

**Q:** Do the compounds follow **Lipinski’s rule of five**? Put a ‘Yes’ or ‘No’ in a new column of your Excel table

**SHOW GOLD POSES FOR THESE 6 LIGANDS:**

* Show the docking poses in the active site of your enzyme in a similar fashion as the VDS3 lab (define active site, hide rest of protein, show hydrophobics as sticks, label 4-5 residues). **Zoom in sufficiently as well. Give your figures super-awesome captions!**

**Polar Contacts**

Find **polar contacts** in PyMol and list them in a new column in the Excel table.

**Q:** Do the ligands dock well into the receptor – i.e. are the scores good or bad relative to what we have seen?

**Q:** Does the random control score poorly?

**Q:** Do the positive controls score well?

**Q:** Which positive control is the best one to use to test this system?

**Q:** How different is the GOLD docking pose of the original ligand relative to the pose from the crystal structure? Did GOLD do a good job or re-placing the ligand in the actual pose?

**Q:** How could we improve this validation dock?